Runchu Tian

506 S. Wright St., Urbana, IL 61801
Tel: (+1) 447-902-6836
Email: runchutian@gmail.com

EDUCATION

University of Illinois Urbana-Champaign GPA: 4.00/4.00
Master of Science in Computer Science Sep. 2024 - Present

Tsinghua University GPA: 3.85/4.00
Bachelor of Science in Mathematics and Physics Sep. 2020 - Jun. 2024

RESEARCH EXPERIENCE

University of Illinois Urbana-Champaign, Data Mining Group Sep. 2024 - Present
Adviser: Jiawei Han, Michael Aiken Chair Professor, Siebel School of Computing and Data Science

Title: *Beyond True or False: Retrieval-Augmented Hierarchical Analysis of Nuanced Claims* ACL 2025
Authors: Priyanka Kargupta*, **Runchu Tian***, Jiawei Han
Overview: Developed ClaimSpect, a retrieval-augmented framework that hierarchically deconstructs nuanced claims into structured aspect trees enriched with corpus-specific evidence and perspectives for in-depth analysis. Oct. 2024 - Feb. 2025

Title: *Retrieval and Structuring Augmented Generation with Large Language Models* KDD 2025
Authors: Pengcheng Jiang, Siru Ouyang, Yizhu Jiao, Ming Zhong, **Runchu Tian**, Jiawei Han
Overview: Reviewed the critical role of structured knowledge in RAG. Nov. 2024 - Jun. 2025

Title: *Deepretrieval: Hacking real search engines and retrievers with large language models via reinforcement learning* COLM 2025
Authors: Pengcheng Jiang, Jiacheng Lin, Lang Cao, **Runchu Tian**, SeongKu Kang, Zifeng Wang, Jimeng Sun, Jiawei Han
Overview: Explored GRPO-based methods for optimizing LLM query rewriting for search engines. Dec. 2024 - Jul. 2025

Title: *Topic Coverage-based Demonstration Retrieval for In-Context Learning* EMNLP 2025
Authors: Wonbin Kweon, SeongKu Kang, **Runchu Tian**, Pengcheng Jiang, Jiawei Han, Hwanjo Yu
Overview: Proposed TopicK, a topic coverage-based retrieval framework that selects demonstrations by modeling the fine-grained topical knowledge needs for effective in-context learning. Feb. 2025 - May 2025

Title: *LLM-Based Compact Reranking with Document Features for Scientific Retrieval* Under Review (TMLR)
Authors: Runchu Tian, Xueqiang Xu, Bowen Jin, SeongKu Kang, Jiawei Han
Overview: Proposed a compact form of document representation in LLM-based listwise reranking. May 2025 - Aug. 2025

Title: *PairSem: LLM-Guided Pairwise Semantic Matching for Scientific Document Retrieval* Under Review (WWW 2026)
Authors: Wonbin Kweon, **Runchu Tian**, SeongKu Kang, Pengcheng Jiang, Zhiyong Lu, Jiawei Han, Hwanjo Yu
Overview: Developed PairSem, an unsupervised LLM-guided framework that models scientific concepts as entity–aspect pairs to enable fine-grained, retriever-agnostic scientific document retrieval. May 2025 - Aug. 2025

University of California San Diego, Shang Data Lab (as visiting student) Jun. 2025 - Aug. 2025
Adviser: Jingbo Shang, Associate Professor, Department of Computer Science and Halıcıoğlu Data Science Institute

Title: *Finish First, Perfect Later: Test-Time Token-Level Cross-Validation for Diffusion Large Language Models* Under Review (ICLR 2026)
Authors: Runchu Tian*, Junxia Cui*, Xueqiang Xu, Feng Yao, Jingbo Shang

Overview: Introduced a test-time decoding strategy that enables token-level cross-validation and iterative refinement in diffusion LLMs, significantly improving generation quality across diverse language tasks. Jul. 2025 - Oct. 2025 (continued remotely)

Tsinghua University, Natural Language Processing Lab
Adviser: Zhiyuan Liu, Associate Professor, Department of Computer Science and Technology

Title: *ToolLLM: Facilitating Large Language Models to Master 16000+ Real-world APIs* ICLR 2024 spotlight
Authors: Yujia Qin, Shihao Liang, Yining Ye, Kunlun Zhu, Lan Yan, Yaxi Lu, Yankai Lin, Xin Cong, Xiangru Tang, Bill Qian, Sihan Zhao, Lauren Hong, **Runchu Tian**, Ruobing Xie, Jie Zhou, Mark Gerstein, Dahai Li, Zhiyuan Liu, Maosong Sun
Overview: Proposed a comprehensive framework for enabling large language models to use external APIs through automated dataset construction (ToolBench), a reasoning-enhanced training algorithm, an evaluation suite (ToolEval) and a language model (ToolLLaMA), which achieves ChatGPT-level tool-use performance. Mar. 2023 - Jun. 2023

Title: *DebugBench: Evaluating debugging capability of large language models* ACL 2024
Authors: Runchu Tian*, Yining Ye*, Yujia Qin, Xin Cong, Yankai Lin, Yinxu Pan, Yesai Wu, Haotian Hui, Weichuan Liu, Zhiyuan Liu, Maosong Sun
Overview: Introduced a benchmark of 4,253 instances across multiple programming languages and bug types to systematically evaluate and analyze the debugging capabilities of large language models. Sept. 2023 - Feb. 2024

Title: *Distance between Relevant Information Pieces Causes Bias in Long-Context LLMs* ACL 2025
Authors: Runchu Tian*, Yanghao Li*, Yuepeng Fu, Siyang Deng, Qinyu Luo, Cheng Qian, Shuo Wang, Xin Cong,Zhong Zhang,Yesai Wu,Yankai Lin...
Overview: Introduced a benchmark for evaluating positional bias in large language models when handling multiple pieces of relevant information, revealing the significant bias based on the spacing of relevant information. May 2024 - Aug. 2024

Title: *Tool Learning with Foundation Models* ACM Computing Surveys, 2024
Authors: Yujia Qin, Shengding Hu, Yankai Lin, Weize Chen, Ning Ding, Ganqu Cui, Zheni Zeng, Yufei Huang, Chaojun Xiao, Chi Han, Yi Ren Fung, Yusheng Su, Huadong Wang, Cheng Qian, **Runchu Tian**, Kunlun Zhu, Shihao Liang, Xingyu Shen ...
Overview: Provided a comprehensive survey and framework for tool learning with foundation models. Jan. 2023 - May 2023

Title: *Exploring Format Consistency for Instruction Tuning* TMLR
Authors: Shihao Liang*, **Runchu Tian***, Kunlun Zhu*, Yujia Qin, Huadong Wang, Xin Cong, Zhiyuan Liu, Xiaojiang Liu, Maosong Sun
Overview: Introduced a framework that automatically transfers instruction formats across datasets to ensure format consistency in instruction tuning, showing that consistent formats improve model performance. Jan. 2023 - June 2023

* denotes equal contribution

PROFESSIONAL EXPERIENCE

University of Illinois Urbana-Champaign, Siebel School of Computing and Data Science
Teaching Assistant, CS 173 - Discrete Structures Sept. 2024 – Dec. 2024
Research Assistant, on *Scientific Claim Verification* advised by Prof. Jiawei Han Jan. 2025 – May 2025
Research Assistant, on *Geospatial Understanding Information System* advised by Prof. Jiawei Han May 2025 – Present

HONORS AND AWARDS

Tsinghua University Academic Excellence Scholarship 2021
Tsinghua University Academic Excellence Scholarship 2022